

Procesamiento de Información Temporal
Tema 2:
Modelos Ocultos de Markov (HMM)
Parte 1: Introducción a los HMM

Alicia Lozano Díez

alicia.lozano@uam.es

Audias – Audio, Data Intelligence and Speech

Universidad Autónoma de Madrid

<http://audias.ii.uam.es>

Materiales basados en los de Daniel Ramos Castro

1

Modelos Ocultos de Markov (HMM):
Introducción

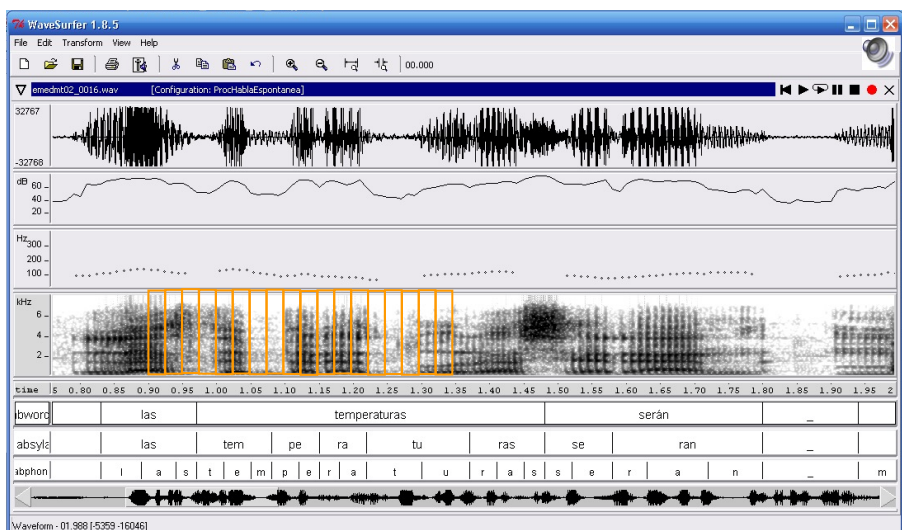
2

Definición Informal de un Hidden Markov Model (HMM)

- Modelo probabilístico
- Basado en un proceso de Markov
- Con estados latentes
 - Nunca se conoce en qué estado estamos
 - Modelo “oculto”
- Del que depende otro proceso “observado”
- Objetivo: determinar el proceso “oculto” a partir de las observaciones

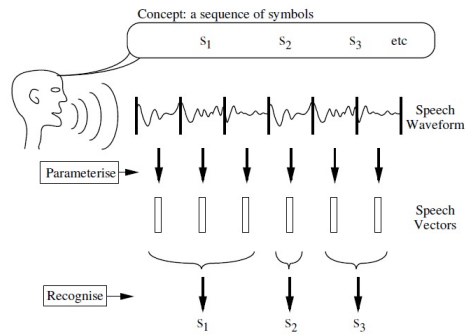
Modelos Ocultos de Markov: Ejemplo

- Ejemplo: Voz
 - Observamos el espectro de voz, pero no sabemos el fonema dicho



Modelos Ocultos de Markov: Ejemplo

- Ejemplo: Voz
 - Estados latentes
 - Por ejemplo, símbolos
 - Observaciones
 - Características extraídas de la voz
 - En ocasiones, a la extracción se le llama "parameterization"
 - Ejemplo: coeficientes del filtro LPC de un fragmento de voz a corto plazo (LPCC)
- Objetivo:
 - A partir de la voz (LPCC)
 - Obtener los símbolos "latentes"



The HTK Book. <https://htk.eng.cam.ac.uk/docs/docs.shtml>

Observaciones: Parametrización

- También llamada "extracción de características"
 - Cada ventana de audio se transforma en un vector de longitud fija
 - Ejemplo: coeficientes de un filtro LPC (con transformada "cepstrum": LP Cepstral Coefficients)
 - Otras variantes más usadas: coeficientes Mel Frequency Cepstral Coefficients (MFCC)
 - Mismo tipo de información que los LPCC, pero suelen funcionar mejor

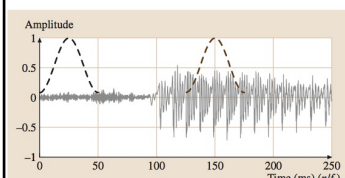
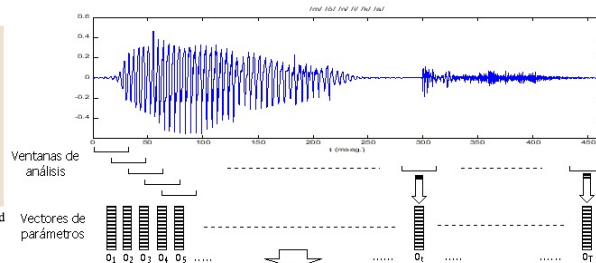
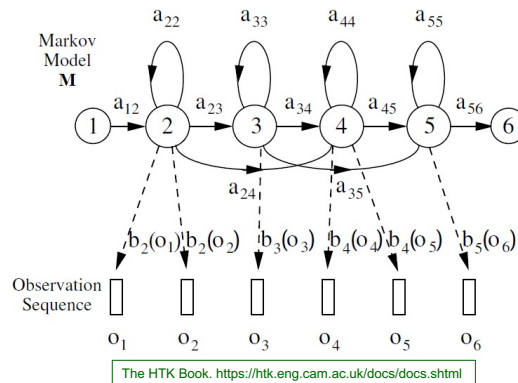


Fig. 9.6 Segment of sampled speech waveform (samples connected by straight lines for plotting)



HMM con Observaciones Continuas: Ejemplo

- Modelo izquierda-derecha
- Observaciones modeladas con un Modelo de Mixturas de Gaussianas (GMM, lo veremos más tarde)



The HTK Book. <https://htk.eng.cam.ac.uk/docs/docs.shtml>

< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 2: Modelos Ocultos de Markov (HMM)

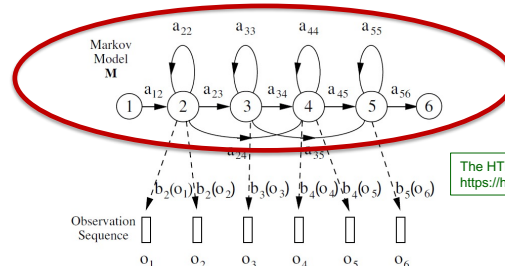
UAM

7

7

HMM con Observaciones Continuas: Ejemplo

- Modelos probabilísticos generativos de estados finitos
 - El modelo está en el estado " i " (círculos con números) en el instante temporal n , con cierta probabilidad
 - En cada momento temporal, el modelo genera una observación de acuerdo con una función densidad de probabilidad $b_i(o)$
 - Caso continuo del ejemplo: $b_i(o)$ es un GMM
 - Podría ser otro tipo de función (categórica, densidad paramétrica, etc.)
 - Tras ello, el modelo se actualiza, pasando al estado j con probabilidad de transición a_{ij}



The HTK Book. <https://htk.eng.cam.ac.uk/docs/docs.shtml>

< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
Tema 2: Modelos Ocultos de Markov (HMM)

UAM

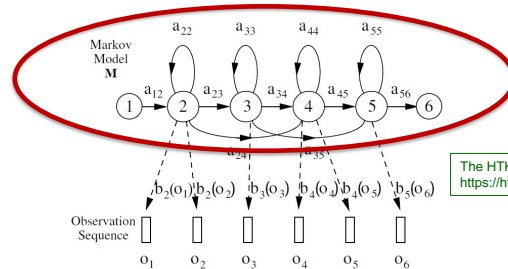
8

8

HMM con Observaciones Continuas: Ejemplo

■ Otros detalles:

- En ocasiones, es útil tener estados que “no generan” (ej. 1, 6)
 - Útil en implementaciones en las que hay que “concatenar” modelos
- HMM Ergódico: todas las probabilidades de transición son >0
- HMM izquierda-derecha, también llamado “Bakis”
 - No se puede ir “atrás” en el grafo: típico para modelar información temporal
 - En voz y audio: estados representan fragmentos de señal
 - Típicamente un fonema se representa con entre 3 y 5 estados



The HTK Book.
<https://htk.eng.cam.ac.uk/docs/docs.shtml>

< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
 Tema 2: Modelos Ocultos de Markov (HMM)

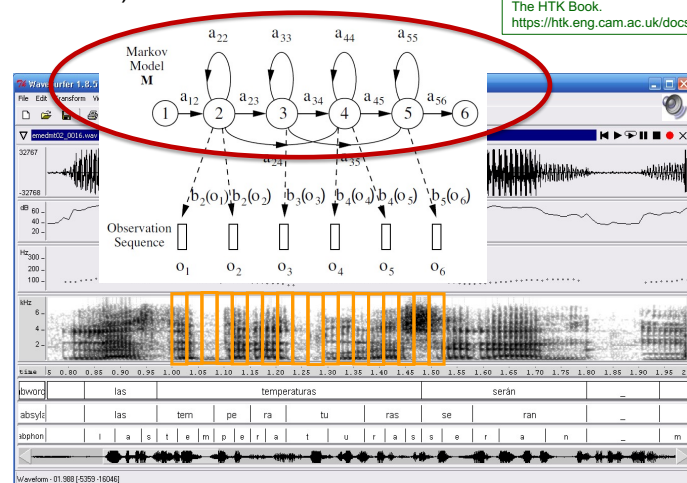
UAM

9

9

HMM con Observaciones Continuas: Ejemplo

- HMMs izquierda-derecha (Bakis): buen modelo para procesar información temporal (como la voz)



The HTK Book.
<https://htk.eng.cam.ac.uk/docs/docs.shtml>

< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
 Tema 2: Modelos Ocultos de Markov (HMM)

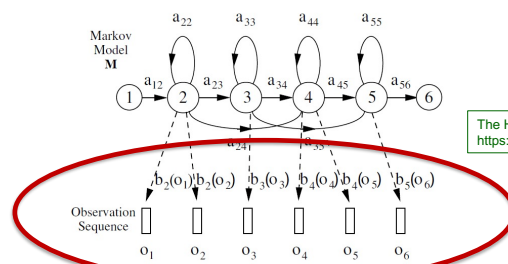
UAM

10

10

HMM-GMM: Gaussian Mixture Models

- Modelo de mixturas de gaussianas (GMM):
 - Función densidad de probabilidad paramétrica (multivariada)
 - Puede utilizarse para generar observaciones de un HMM
 - Ejemplo: Distribución multivariada de las características LPCC o MFCC de cada fragmento de audio o voz
 - El HMM otorga la estructura temporal: una distribución diferentes (asociada a un estado) es más o menos probable para cada momento en el tiempo



The HTK Book.
<https://htk.eng.cam.ac.uk/docs/docs.shtml>

< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
 Tema 2: Modelos Ocultos de Markov (HMM)

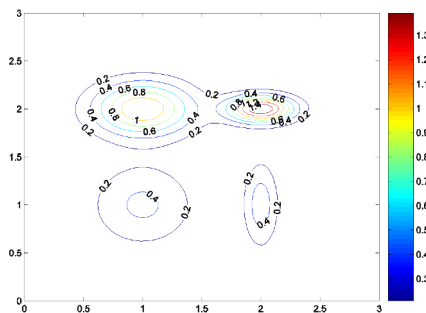
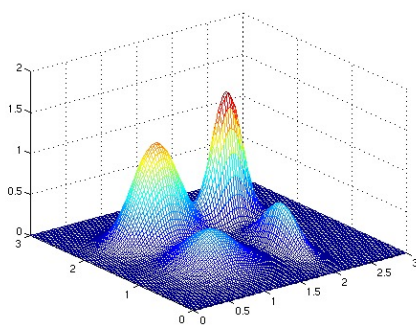
UAM

11

11

HMM-GMM: Gaussian Mixture Models

- Modelos de Mixturas de Gaussianas
 - GMM con 4 mixturas y 2 dimensiones



< audias >

Procesamiento de Información Temporal. Máster en Ciencia de Datos
 Tema 2: Modelos Ocultos de Markov (HMM)

UAM

12

12

HMM-GMM: Gaussian Mixture Models

- GMMs definidos por un conjunto reducido de parámetros

$$p(\mathbf{x}|\lambda_p) = \sum_{i=1}^M \omega_{ip} g_{ip}(\mathbf{x})$$

$$g_{ip}(\mathbf{x}) = N(\mu_{ip}, \Sigma_{ip})$$

x_i : coeficiente MFCC en la trama de voz en el instante temporal "t"

Vectores de medias:

$$\mu_p = \{\mu_{ip}\}$$

Matrices de covarianzas:

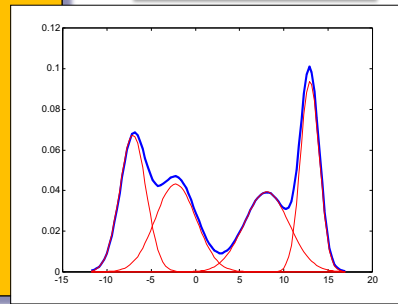
$$\Sigma_p = \{\Sigma_{ip}\}$$

Pesos de cada mixtura:

$$\omega_p = \{\omega_{ip}\}, \sum_i \omega_{ip} = 1$$

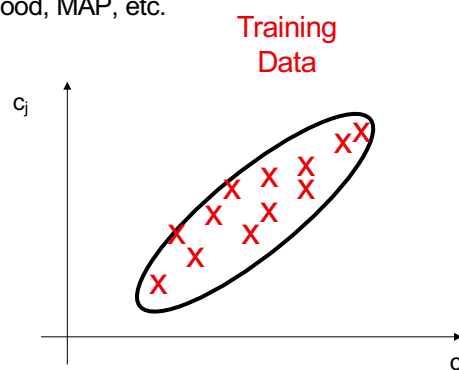
GMM para un estado "p":

$$\lambda_p = \{\mu_{ip}, \Sigma_{ip}, \omega_{ip}\}$$



HMM-GMM: Gaussian Mixture Models

- Una gaussiana fácil aprendizaje si se conocen "sus" observaciones
 - Maximum Likelihood, MAP, etc.



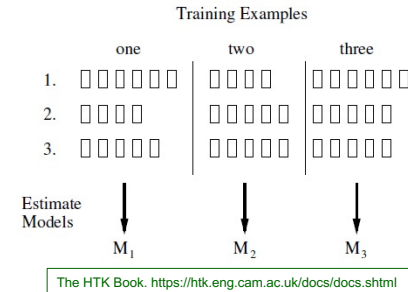
- El problema es aprender la mixtura entera
 - No sabemos qué observación corresponde a cada mixtura
- Veremos una solución más adelante

HMM-GMM: Problemas a Resolver

■ Entrenamiento:

- Tenemos un número de ejemplos (observaciones) etiquetados
- Queremos obtener un buen modelo para esos ejemplos
 - Parámetros del HMM y del GMM

(a) Training



■ Para ello, necesitamos...

- Un método para asociar observaciones a estados del HMM
 - Alineamiento de estados
- Un método eficiente de “puntuar” un modelo con unas observaciones
 - Puntuación

HMM-GMM: Tareas

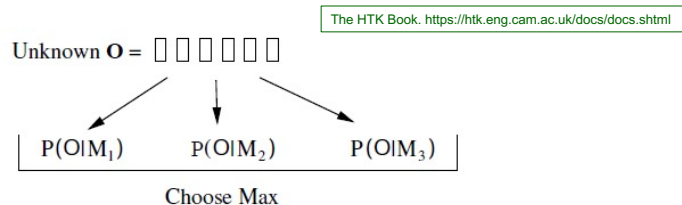
■ Generación de observaciones

- Dado un modelo entrenado, generar observaciones de ese modelo
- Para síntesis, etc.

■ Reconocimiento:

- Dada una secuencia de observaciones no etiquetadas, obtener las etiquetas
 - Por ejemplo, a qué secuencia de fonemas pertenecen unas observaciones de voz

(b) Recognition



HMM-GMM: Ejemplo de Tarea de Reconocimiento

- Reconocimiento de dígitos aislados
 - Voz pronunciando dígitos ("uno", "dos", "tres"....)
- Entrenamiento
 - Se debe disponer de observaciones etiquetadas con el dígito que se pronuncia
 - Se entrena un GMM-HMM por cada dígito
 - Usando todas las observaciones de dicho dígito
- Reconocimiento
 - Se tiene una secuencia de observaciones cuyo dígito no se conoce
 - Se compara dicha secuencia con todos los HMM entrenados
 - Se elige el que otorgue la puntuación mayor: ese HMM determina el dígito pronunciado

(b) Recognition

Unknown $O = \square \square \square \square \square$

The HTK Book. <https://htk.eng.cam.ac.uk/docs/docs.shtml>

$P(O|M_1)$ $P(O|M_2)$ $P(O|M_3)$

Choose Max

Referencias Bibliográficas

- [Young02] S. Young et al. The HTK Book. 2002. Disponible bajo registro en <https://htk.eng.cam.ac.uk/docs/docs.shtml>.
- [Rabiner89] L. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. Proceedings of the IEEE, 77(2), 1989.